

## Solutions to Problem Set 4

4-1 Suppose you are given an  $m \times n$  matrix  $A \in \mathbb{R}^{m \times n}$  with row sums  $r_1, \dots, r_m \in \mathbb{Z}$  and column sums  $c_1, \dots, c_n \in \mathbb{Z}$ . Some of the entries might not be integral but the row sums and column sums are. Show that there exists a rounded matrix  $A'$  with the following properties:

- row sums and column sums of  $A$  and  $A'$  are identical,
- $a'_{ij} = \lceil a_{ij} \rceil$  or  $a'_{ij} = \lfloor a_{ij} \rfloor$  (i.e.  $a'_{ij}$  is  $a_{ij}$  either rounded up or down).

By the way, this rounding is useful to the census bureau as they do not want to publish statistics that would give too much information on specific individuals. They want to be able to modify the entries without modifying row and column sums.

Let  $G = (V, E)$  be a digraph with  $m + n + 2$  vertices constructed as following:

- The set of vertices  $V$  is consist of the source  $s$ , the sink  $t$ , the set of  $m$  vertices  $A = \{a_1, \dots, a_m\}$ , and the set of  $n$  vertices  $B = \{b_1, \dots, b_n\}$ .
- The set of edges  $E$  is defined as

$$E = \{(s, a_i) : i \in [m]\} \cup \{(a_i, b_j) : i \in [m], j \in [n]\} \cup \{(b_j, t) : j \in [n]\}.$$

Let  $u : E \rightarrow \mathbb{R}$  be the upper capacity function defined as

$$u(e) = \begin{cases} r_i & \text{if } e = (s, a_i) \\ s_j & \text{if } e = (b_j, t) \\ \lceil a_{ij} \rceil & \text{if } e = (a_i, b_j), \end{cases}$$

and  $\ell : E \rightarrow \mathbb{R}$  be the lower capacity function defined as

$$\ell(e) = \begin{cases} r_i & \text{if } e = (s, a_i) \\ s_j & \text{if } e = (b_j, t) \\ \lfloor a_{ij} \rfloor & \text{if } e = (a_i, b_j). \end{cases}$$

Let  $x : E \rightarrow \mathbb{R}$  be the assignment of values to the arcs so that

$$x(e) = \begin{cases} r_i & \text{if } e = (s, a_i) \\ s_j & \text{if } e = (b_j, t) \\ a_{ij} & \text{if } e = (a_i, b_j). \end{cases}$$

Clearly,  $\ell(e) \leq x(e) \leq u(e)$  for any  $e \in E$  and  $x$  satisfies all flow conservation constraints. Hence, the flow problem is feasible. By Corollary 4.2 (of lecture note), there exists a flow  $x'$  with integral values, i.e.,  $x'(e) \in \mathbb{Z}$  for all  $e \in E$ .

Let  $A'$  be the matrix in  $\mathbb{Z}^{m \times n}$  with entries  $a'_{ij} = x'((a_i, b_j))$ . Then, by flow conservation constraints, row sums of  $A'$  are  $r_1, \dots, r_m$  and column sums of  $A'$  are  $s_1, \dots, s_n$  as desired.

- 4-2 At some point during baseball season, each of  $n$  teams of the American League has already played several games. Suppose team  $i$  has won  $w_i$  games so far, and  $g_{ij} = g_{ji}$  is the number of games that teams  $i$  and  $j$  have yet to play. No game ends in a tie, so each game gives one point to either team and 0 to the other. You would like to decide if your favorite team, say team  $n$ , can still win. In other words, you would like to determine whether there exists an outcome to the games to be played (remember, with no ties) such that team  $n$  has at least as many victories as all the other teams (we allow team  $n$  to be tied for first place with other teams).

Show that this problem can be solved as a maximum flow problem. Give a necessary and sufficient condition on the  $g_{ij}$ 's so that team  $n$  can still win.

Notice that one can begin by checking whether  $w_i > w_n + g'$ , where  $g' := \sum_{i=1}^{n-1} g_{ni}$ , for some  $i < n$ . Obviously if this was the case, team  $n$  could not win. Therefore, let us assume  $w_i \leq w_n + \sum_{i=1}^{n-1} g_{ni}$  for all  $i < n$ . Moreover, if there was an outcome where team  $n$  won, then it could only get better if it won all of its games, so let us assume that all  $g_{ni}$  games to be played between teams  $n$  and  $i$  are won by team  $n$ . If we let  $x_{ij}$  to be the number of games between  $i$  and  $j$  won by  $i$ , then team  $n$  has a chance of winning iff there are positive integers  $x_{ij}$  with  $x_{ij} + x_{ji} = g_{ij} = g_{ji}$ ,  $i \neq j$ , and  $w_i + \sum_{j=1}^{n-1} x_{ij} \leq w_n + g'$  for all  $i < n$ .

Consider the graph  $G$  on  $V = \{s, t\} \cup \{v(i, j)\}_{1 \leq i < j \leq n-1} \cup \{w(i)\}_{1 \leq i \leq n-1}$  with edges classified in these categories (all lower capacities are  $l(e) = 0$ ).

- All edges  $e$  from  $s$  to  $v(i, j)$  with  $u(e) = g_{ij}$ .
- All edges  $e_1, e_2$  from  $v(i, j)$  to  $w(i)$  and to  $w(j)$  with  $u(e_1) = u(e_2) = \infty$ .
- All edges  $e$  from  $w(i)$  to  $t$  with  $u(e) = w_n + g' - w_i$ .

We claim that team  $n$  can win iff the maximum flow from  $s$  to  $t$  is  $\sum_{i < j < n} g_{ij}$ . Indeed a maximum flow with that value exists iff there exists an integer flow  $x : E \rightarrow \mathbb{Z}$  with that flow value exists (since all capacities are integers). If we let  $y_{ij} = x(v(i, j), w(i))$ , such integer flow satisfies  $y_{ij} \geq 0$ ,  $g_{ij} = y_{ij} + y_{ji}$  and  $\sum_{j=1}^{n-1} x_{ij} \leq w_n + g' - w_i$ . As asserted above, this is equivalent to team  $n$  having some chance.

**Remark.** Note that we can use the max-flow/min-cut theorem to give an alternate characterization of when team  $n$  can win. Suppose team  $n$  can't win, by the above, this is equivalent to the maximum flow in the graph  $G$  being bounded strictly below  $\sum_{i < j < n} g_{ij}$ . By max-flow/min-cut, this is equivalent to the existence of a cut  $\delta^+(S)$  with capacity less than  $\sum_{i < j < n} g_{ij}$ ,  $S \subseteq V$ ,  $s \in S$ ,  $t \notin S$ . Let  $W$  be the set of  $i$  such

that  $w(i)$  is in  $S$ , then in order for the capacity of  $\delta^+(S)$  to be below  $\sum_{i < j < n} g_{ij}$ , we must have

$$\sum_{\substack{i, j \in W \\ i < j}} g_{ij} > \sum_{i \in W} w_n + g' - w_i.$$

So team  $n$  has a chance of winning if and only if a set  $W$  satisfying the above inequality does not exist.

4-3 Consider the following orientation problem. We are given an undirected graph  $G = (V, E)$  and integer values  $p(v)$  for every vertex  $v \in V$ . We would like to know if we can orient the edges of  $G$  such that the directed graph we obtain has at most  $p(v)$  arcs incoming to  $v$  (the “indegree requirements”). In other words, for each edge  $\{u, v\}$ , we have to decide whether to orient it as  $(u, v)$  or as  $(v, u)$ , and we would like at most  $p(v)$  arcs oriented towards  $v$ .

1. Show that the problem can be formulated as a maximum flow problem. That is, show how to create a maximum flow problem such that, from its solution, you can decide whether or not the graph can be oriented and if so, it also gives the orientation.
  2. Consider the case that the graph cannot be oriented and meet the indegree requirements. Prove from the max-flow min-cut theorem that there must exist a set  $S \subseteq V$  such that  $|E(S)| > \sum_{v \in S} p(v)$ , where as usual  $E(S)$  denotes the set of edges with both endpoints within  $S$ .
1. Let  $G$  be the given undirected graph. We construct a directed graph  $D$  from  $G$  as follows:
    - For each edge  $e = \{v_1, v_2\} \in E(G)$ , we add a new vertex  $w_e$  and replace  $e$  by two directed edges  $(w_e, v_1)$  and  $(w_e, v_2)$ .
    - We add the source  $s$  and connect it to  $w_e$  for all  $e \in E(G)$ .
    - We add the sink  $t$  and connect every vertex  $v \in V(G)$  to  $t$ .

We let the lower capacity be 0, and the upper capacity be  $u : E(D) \rightarrow \mathbb{R}$  where

$$u(a) = \begin{cases} p(v) & \text{if } a = (v, t) \text{ for } v \in V(G) \\ 1 & \text{if } a = (s, w_e) \text{ for } e \in E(G) \\ +\infty & \text{otherwise.} \end{cases}$$

**Claim.** There exists an orientation meeting the in-degree requirement if and only if the value of a maximum-flow is equal to  $|E(G)|$ .

**Proof:** Given any orientation of  $E(G)$ , let us define a valuation  $x : E(D) \rightarrow \mathbb{Z}$  so that

- for any edge  $e = \{u, v\} \in E(G)$  we have  $x((w_e, u)) = 1$  and  $x((w_e, v)) = 0$  if  $e$  is oriented as  $(v, u)$  or  $x((w_e, u)) = 0$  and  $x((w_e, v)) = 1$  if  $e$  is oriented as  $(u, v)$ ,
- $x((s, w_e)) = 1$  for all  $e \in E(G)$ , and
- $x((v, t)) = \sum_{e \in E(G): v \in e} x((w_e, v))$  for all  $v \in V(G)$ .

Note that  $x$  satisfies flow conservation constraints by definition. Furthermore,  $x((v, t))$  is exactly the in-degree of  $v$  in the orientation. It implies that if the given orientation meets the in-degree requirement, then  $x$  is a feasible flow with the value  $|E(G)|$ , which is maximum.

To prove the converse, let  $x : E(D) \rightarrow \mathbb{Z}$  be an integral flow of value  $|E(G)|$ . Note that

$$|x| = \sum_{e \in E(G)} x((s, w_e)) \leq \sum_{e \in E(G)} u((s, w_e)) = |E(G)|,$$

so we must have  $x((s, w_e)) = 1$  for all  $e \in E(G)$ . Furthermore, since  $x$  is integral, we have  $x((w_e, u))$  or  $x((w_e, v))$  equal to 1 where  $u$  and  $v$  are endpoints of  $e$ . Now orient  $e$  as  $(u, v)$  if  $x((w_e, v)) = 1$  or  $(v, u)$  if  $x((w_e, u)) = 1$ . In this orientation the in-degree of a vertex  $v \in V(G)$  is equal to

$$\sum_{e: v \in e} x((w_e, v)) = x((v, t)),$$

which is at most  $u((v, t)) = p(v)$  as desired.

△

2. By max-flow min-cut theorem, there is a flow of value  $|E(G)|$  if and only if

$$|E(G)| \leq \sum_{a \in \delta^+(S')} u(a)$$

for any  $S' \subseteq V(D)$  such that  $s \in S'$  but  $t \notin S'$ .

Suppose that  $G$  cannot be oriented to meet the in-degree requirements. In the first part, we proved that there is no flow of value  $|E(G)|$ , so there is  $S' \subseteq V(D)$  with  $s \in S'$  and  $t \notin S'$  such that

$$C(S') = \sum_{a \in \delta^+(S')} u(a)$$

is minimum and  $C(S') < |E(G)|$ . Since the edges of the form  $(w_e, v)$  have infinite capacities, they do not participate in  $\delta^+(S')$ . Therefore, vertex  $w_e$  is present in  $S'$  if and only if both endpoints of  $e$  are present in  $S'$ . In other words, if we set  $S = S' \cap V(G)$ , then

$$S' = \{s\} \cup \{w_e : e \in E(S)\} \cup S.$$

Hence,

$$\begin{aligned} C(S') &= \sum_{e \notin E(S)} u((s, w_e)) + \sum_{v \in S} u((v, t)) \\ &= |E(G)| - |E(S)| + \sum_{v \in S} p(v) < |E(G)| \end{aligned}$$

so we have

$$\sum_{v \in S} p(v) < |E(S)|.$$

4-5 Prove that the cut function of an undirected graph is submodular.

Let  $f(S) = u(\delta(S))$  be the cut function in an undirected graph with capacities. We need to show that

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B).$$

Consider an arbitrary edge  $e$ . Below is the table of what  $u(e)$  contributes to both sides of this inequality. The rows and the columns indicate where the endpoints of  $e$  lie.

	$A \setminus B$	$B \setminus A$	$A \cap B$	$V \setminus (A \cup B)$
$A \setminus B$	$0 \geq 0$	$2 \geq 0$	$1 \geq 1$	$1 \geq 1$
$B \setminus A$	$2 \geq 0$	$0 \geq 0$	$1 \geq 1$	$1 \geq 1$
$A \cap B$	$1 \geq 1$	$1 \geq 1$	$0 \geq 0$	$2 \geq 2$
$V \setminus (A \cup B)$	$1 \geq 1$	$1 \geq 1$	$2 \geq 2$	$0 \geq 0$

Adding up those inequalities over all edges, we get the result.

4-6 Show that the definition of submodularity is equivalent to the property of *having diminishing returns*. A function  $f : 2^V \rightarrow \mathbb{R}$  has diminishing returns if for all  $S \subseteq T$  and  $v \notin T$  we have

$$f(T \cup \{v\}) - f(T) \leq f(S \cup \{v\}) - f(S).$$

One direction is easy: if  $f$  is submodular, and if  $S \subseteq T$ ,  $v \notin T$ , then apply the submodularity with  $A = T$ ,  $B = S \cup \{v\}$ :

$$f(T) + f(S \cup \{v\}) \geq f(T \cup \{v\}) + f(S),$$

so  $f$  has diminishing returns.

For the other direction, suppose  $f$  has diminishing returns, and take arbitrary  $A, B \subseteq V$ . We induct on  $|B \setminus A|$ . If  $B \subseteq A$  then there is nothing to prove. Now, let  $|B \setminus A| \geq 1$ . Let  $v \in B \setminus A$  be any element, and let  $B' = B \setminus \{v\}$ . We have:

$$\begin{aligned} f(A) - f(A \cap B) &= f(A) - f(A \cap B') \stackrel{\text{ind. hyp.}}{\geq} f(A \cup B') - f(B') \\ &\stackrel{\text{dim. ret.}}{\geq} f(A \cup B' \cup \{v\}) - f(B' \cup \{v\}) = f(A \cup B) - f(B) \end{aligned}$$

completing the proof.